HUMAN CANCER-RELATED GENE, ITS ENCODED PRODUCTS AND APPLICATIONS

FIELD OF THE INVENTION

The present invention relates to a human cancer-related gene, its encoded products and their applications in genetic engineering and protein engineering areas, as well as in medical diagnosis and treatment.

BACKGROUND OF THE INVENTION

Cancer is the major health problem threatening human lives. Hepatocellular carcinoma (HCC) is one of the most serious among cancer diseases. It is reported that the new cases of primary hepatocellular carcinoma exceeds over one million worldwide each year. 70% of the new cases occur in Asia, and about 40-45% of the worldwide new cases occur in China. The total number of new hepatocellular carcinoma cases every year in China is about 450,000, and the number is increasing, especially in those between ages 20-60. The high incidence, difficulty in early diagnosis, fast growing rate, high reoccurrence, and the high mortality rate make HCC a most malignant cancer. Most HCC patients have already progressed to the intermediate stage or late stage when diagnosed, and they can only survive for 3-6 months if without a proper treatment.

To elucidate the mechanism underlying cancerogenesis would help for cancer prevention, diagnosis and treatment. Early diagnosis is crucial for raising the curative rate and reducing the mortality. Currently used HCC-diagnostic marker, the serum AFP, has 30% of negative results in HCC patients, while some benign liver disease can cause a significant increase of AFP level in serum, creating some difficulty in differential diagnosis. It has been found that the hepatocarcinogenesis is related to individual hereditary susceptibility. Individuals with different genetic backgrounds possess different handling capability toward environmental carcinogens. This leads to different risk of suffering from cancer for individuals. It is the various genotypes and the genetic diversity that cause the different genetic susceptibility for cancerogenesis.

Cancer is essentially a cellular hereditary disease. Although a great number of cancer-related genes have been discovered, the mechanisms of the cancerogenesis and the

development remain to be elucidated. The known oncogenes can be divided into five categories according to the cellular localization and function of their encoded proteins: I. genes that encode growth factors, including sis, int-2, hst, fgf-5; II. genes that encode growth factor receptors, including erbB, erbB-2, fms, met, ros, and others; III. genes that encode signal transduction molecules in cytoplasm, including abl, src, ras, raf, yes, fgr, fes, lck, mos, and others; IV. genes that encode regulatory molecules for cell proliferation and apoptosis, including bcl-1, bcl-2 and others; and V. genes that encode the nuclear DNA-binding proteins (transcription factors) , such as myc, myb, fos, jun, B-lym, ski, ets, rel and others. It has been demonstrated that ras, src, myc, met and p53 etc. are the genes closely associated with HCC.

SUMMARY OF THE INVENTION

This invention provides a novel human cancer-related gene and its encoded products.

This novel human cancer-related gene provided by this invention is designated as *LAPTM4B*. It comprises one of the following nucleotide sequences:

1. The human cancer-related gene comprises one of the following nucleotide sequences:

1). SEQ ID No: 1, SEQ ID No: 2, SEQ ID No: 3 or SEQ ID No: 6 in the sequence listings;

2). Polynucleotides that encode the protein sequences of SEQ ID No: 4, SEQ ID No: 5, or SEQ ID No: 7 in the sequence listings;

3). DNA sequences having more than 90% homology with the DNA sequences defined by SEQ ID No: 1, SEQ ID No: 2, SEQ ID No: 3 or SEQ ID No: 6 in the sequence listings. These DNA sequences can encode proteins having the same or similar functions.

SEQ ID No: 1 in the sequence listings mentioned above contains 954 bases. It is an intact open reading frame. SEQ ID No: 1 has two starting sites, one is the base at 1-3 site at 5' terminal, and another is the base at 274-276 site at 5' terminal. Two complete cDNAs in SEQ ID No: 1 have two alternative tailing signals. When 5' terminal in SEQ ID No: 1 is extended outward by 85 bases, and 3' terminal is extended outward by 401 bases, SEQ ID No: 2 in the sequence listings is obtained. This gene contains 1440 bases. When 5' terminal in SEQ ID No: 1 is extended outward by 85 bases, and 3' terminal is extended outward 1130 bases, SEQ ID No: 3

in the sequence listings is obtained. This gene is consisted of 2169 bases. *LAPTM4B* gene localizes on chromosome 8q22.1.

In the sequence listings, SEQ ID No: 6 is the allelic gene of SEQ ID No: 1, consisting of 2264 bases. Its open reading frame starts from 7 to 1129 base. This sequence contains two tandemly arranged 19bp DNA segments, the sequence of which is gcttgg agctccagca gct. These 19bp DNA segments localized in nt 124 – nt 161 in SEQ ID No: 6.

The human cancer related LAPTM4B protein possesses the amino acid sequence of 4, and/or 5, and/or 7. Or it consists of the sequence 4, and/or the sequence 5, and/or the sequence 7 after one or several amino acid residues are replaced, deleted, or added. However, the above altered sequence 4, and/or the sequence 5, and/or the sequence 7 still have the same or similar activity to the unchanged sequence 4, and/or the sequence 5, and/or the sequence 7.

Sequence 4 in the sequence listings consists of 317 amino acid residues encoded by the whole sequence of SEQ ID No: 1. Its molecular mass is 35kDa and the putative isoelectric point is 9.05. Sequence 5 in the sequence listings contains 226 amino acid residues encoded by the segment of bases from 274th to 954th in the SEQ ID No: 1. Its molecular mass is 24kDa, and the putative isoelectric point is 4.65. The sequence 7 in the sequence listings is a protein containing 370 amino acid residues.

LAPTM4B gene is widely expressed at different levels in sixteen normal tissues. Its transcriptive expression is very high in testis, cardiac muscle, and skeletal muscle, moderate in ovary, kidney, and pancreas, low in liver, spleen, small intestine, large intestine, and thymus, and is very low in lung and peripheral blood cells. In eight fetal tissues, the expression is high in heart, skeletal muscle, and kidney. In fetal livers, it is slightly higher than that in adult livers. However, its expression in some cancerous tissues is significantly upregulated. For instance, the Northern Blot analysis indicates that the transcription level in 87.3% (48/55) human hepatocellular carcinoma tissues is significantly higher than that in fetal livers and normal livers (Figure 1-A). In situ hybridization (Figure 2-A), immuno-histochemistry (Figure 2-B), and immuno-cytochemistry (Figure 2-C) also indicate that *LAPTM4B* gene expression

is especially high in hepatocellular carcinoma tissues, while its expression is relatively low in paired non-cancerous liver tissues (Figures 2-A and 2-B). Among the five cell lines from hepatoma tissues tested, all except for HLE, SMMC-7721, QGY-7701, BEL7402 and HG116 are expressed highly (Figure 1-B and Figure 2-C). It is important that highly over expressed protein product in hepatocellular carcinoma tissue and hepatocellular carcinoma cell line is mainly SEQ ID No: 4 LAPTM4B-35, while SEQ ID No: 5 LAPTM4B-24 only shows a slightly up regulation in its expression level. This results in a remarkable increase in the ratio of LAPTM4B-35 to LAPTM4B-24 proteins in the hepatocellular carcinoma tissue (Figure 2-B). Although the expressions of LAPTM4B-35 and LAPTM4B-24 are slightly increased in the paired non-cancerous tissue, their ratio is the same as that in the normal liver (See Table 1). This is probably a precancerous sign of hepatocellular carcinoma. In addition, the expression levels of mRNA and the protein of *LAPTM4B* gene is negatively correlated with the differentiation of the hepatocellular carcinoma tissue. The hepatocellular carcinoma tissues in low differentiation are expressed highly, while the ones in high differentiation are expressed relatively low (Figure 1-C). The Western Blot and the immuno-histochemical method are used to determine the relationship of *LAPTM4B* gene with other cancers. The results indicate that LAPTM4B-35 protein expression is up regulated in some epithelium derived cancerous tissues and cell lines, such as stomach cancer, breast cancer, highly metastatic human lung cancer, and prostate cancer (Figure 11). Moreover, in syngeneic human lung cancer and prostate cancer cell lines, LAPTM4B-35 expression is greatly up regulated in cells of high metastasis potential compared with those of low metastasis potential. But in cell lines of human melanoma, either from in situ or metastatic cancer, it is not clearly expressed. Although LAPTM4B-35 is expressed at a low level in liver tissues of adult rats and mice, its expressionis not obviously up regulated in either mouse ascetically qrown hepatocellular carcinoma or in the regenerated rat liver under a normal proliferation status.

Table 1: Expression ratio of LAPTM4B-35 to LAPTM4B-24 in hepatocellular carcinoma tissue, paired non-cancerous liver tissue and normal liver tissue

	HCC	PNL	NL
LAPTM4B-35	13.32±1.98	4.58±1.31	2.78±0.11
LAPTM4B-24	3.59±1.78	1.76±1.24	1.00±0.02
LAPTM4B-35/ LAPTM4B-24 (Ratio)	3.71	2.60	2.78

$P<0.01$ HCC *vs.* PNL and NL

LAPTM4B proteins in SEQ ID No: 4, SEQ ID No: 5, and SEQ ID No: 7 have four fragments of membrane-spanning sequences, one N-glycosylation site, a typical lysosomeand endosome targeting signals in the cytoplasmic region. They all belong to the protein superfamily of the tetratransmembrane proteins. However, they have various number of phosphorylation sites. The experiment shows that SEQ ID No: 4 LAPTM4B-35 can form a complex in plasma membrane with the integrin $\alpha 6\beta 1$ (Single specific receptor of laminin in the extracellular matrix) and the epidermal growth factor recceptor (EGFR) (Figures 14-A, B, and C). This complex is colocalized in cell plasma membrane. It is possible that LAPTM4B-35 may integrate in the plasma membrane the proliferation signals from both extracellular matrix and the growth factor. This can further elucidate molecular mechanism of the anchorage-dependent cell growth of normal eukaryotic cells, i.e. the eukaryotic cell growth needs not only the stimulating signal from the growth factor, but also a definite stimulating signal from extracellular matrix. It represents a break through in understanding the regulation mechanism of the cell proliferation. Experiments demonstrate that tyrosine group (Tyr_{285}) in the cytoplasmic region of LAPTM4B protein C terminal can be phosphorylated (Figure 15-A). When the cell is attaching onto the laminin substrate, its phosphorylation is increased sharply (Figure 15-A) and can be completely inhibited by LAPTM4B-EC2-pAb antibody (Figure 15-B), but the non-correlated antibody does not show any inhibitory effect (Figure 15-C). After the phosphorylation, Tyr_{285} forms a site to bind with the SH2 domain of intracellular signal molecules. In the meantime, N terminal and C terminal sequences of LAPTM4B contain Pro-rich domains and binding sites of the typical SH3 1 domain. The above results indicate that SEQ ID No: 4 LAPTM4B-35 protein may be an important docking protein for signal transduction, or an organizer of the special

microdomain in the plasma membrane. It can recruit related signal molecules from both inside and outside of the cells to complete the signal transduction for cell proliferation, differentiation, and apoptosis. Experimental results show that the transfection of mouse NIH3T3 cells and HLE human hepatocellular carcinoma cells by cDNA in SEQ ID No: 4 produces stable transfected and LAPTM4B-35 over expressed NIH3T3-AE and HLE-AE cell lines. The growth curves (Figure 4), the incorporation of 3H-TdR (Figure 5), and the cell numbers in S phase of cell cycle (Figure 6) all demonstrate that the rate of cell proliferation is greatly increased. Moreover, the proliferation of transfected cells shows less dependence on the growth factor in serum, and the transfected cells can form large colonies in soft agar. Inoculation of NIH mouse with NIH3T3-AE cells can form a moderate malignant fibrosarcoma (Figure 7), indicating the over expression of LAPTM4B-35 induces outof -control of the cell proliferation. Also migration capability of the HLE-AE cells is strengthened and its capability to invade the Matrigel is remarkably enhanced, indicating that the over expression of LAPTM4B-35 accelerates the development of cell malignant phenotype. On the contrary, the cDNA of SEQ ID No: 5 (An encoding sequence where 91 amino acids in the N terminal of LAPTM4B-35 is truncated) transfected mouse BHK, NIH3T3, and HLE hepatocellular carcinoma cell lines cannot survive for a long time. The result shows that the 91 amino acid sequence on the N terminal of SEQ ID No: 4 LAPTM4B-35 protein play a crucial role in regulating cell proliferation. LAPTM4B-35 protein and LAPTM4B-24 protein have reciprocal, antagonistical functions in cell proliferation and survival. The overexpression of LAPTM4B-35 accelerates cellular malignant transformation, while the overexpression of LAPTM4B-24 facilitates the cell death. Their expression equilibration and regulation are pivotal to the carcinogenesis and progression of malignant cancer. LAPTM4B gene may belong to the proto-oncogene family. In cancer treatment, inhibiting SEQ ID No: 4 LAPTM4B-35 expression and strengthening SEQ ID No: 5 LAPTM4B-24 expression may suppress the growth of hepatocellular carcinoma and reverse its malignant phenotype or progressively slow down its development. Furthermore, the overexpression of LAPTM4B-35 also promotes upregulation of the cell cycle regulators, such as cyclin D1 (Figure 13-A) and cyclin E (Figure 13-B), and also the over expression of some proto-oncogenes, such as c-Myc (Figure 13-C), c-Jun (Figure 13-D), and c-Fos (Figure 13-E) etc.

The monoclonal and polyclonal antibodies for SEQ ID No: 4 LAPTM4B-35 protein epitopes, such as polyclonal LAPTM4B-$EC2_{232-241}$-pAb for SEQ ID No: 4 LAPTM4B-35 in the secondary extracellular region, polyclonal antibodies (LAPTM4B-N_{1-99}-pAb and LAPTM4B-N_{28-37}-pAb) for SEQ ID No: 4 LAPTM4B-35 N terminal sequence, and monoclonal antibodies for LAPTM4B are important in studying the effects of LAPTM4B-35 and LAPTM4B-24 in cancer diagnosis and treatment (Figures 2, 3, 8, 11, 12, 14, 15). For example, LAPTM4B-$EC2_{232-241}$-pAb, LAPTM4B-N_{1-99}-pAb polyclonal antibodies and LAPTM4B-N_{1-99}-mAb monoclonal antibody can be used to analyze LAPTM4B protein expression, intracellular localization, separation and purification, and protein-protein interaction. They can also be used to detect the antibody and antigen of LAPTM4B in blood (Figure 8). Moreover, LAPTM4B-$EC2_{232-241}$-pAb can inhibit cancer cell proliferation (Figure 12), Tyr_{285} phosphorylation of LAPTM4B protein (Figure 15-B), and the phosphorylation and activation of signal molecules FAK (Figure 16-A) and MAPK (Figure 16-B). Therefore, all the monoclonal and polyclonal antibodies for SEQ ID No: 4 LAPTM4B-35 protein epitope are encompassed in this invention.

SEQ ID No: 8 is the promoter sequence of *LAPTM4B* gene. To study the regulation of *LAPTM4B* gene expression, the *LAPTM4B* gene promoter and the upstream sequence SEQ ID No: 8 are cloned. There are no typical CCAAT (TTGCGCAAT), TATA cassettes in *LAPTM4B* gene promoter region. But various binding sites of transcription factors exist in the upstream region of *LAPTM4B* promoter, such as CREBP1/c-Jun, CEBP, PAX2/5/8, GATA, STAT, c-Ets-1, E2F, LYF-1, and c/v-Myb (Figure 17A). These transcription factors may specifically regulate *LAPTM4B* expression in cells of various tissues. The abnormal expression and activation of these transcription factors in cancer cells possibly lead to an unbalanced expression of LAPTM4B proteins. Moreover, there are two highly homologous repeating sequences in the upstream domain of *LAPTM4B* promoter. It is worthwhile to study whether they are responsible to the regulation of *LAPTM4B* expression. A series of vectors consisting upstream region sequences of *LAPTM4B* promoter with different lengths—promoter—5' UTR-35bp encoding region—luciferase reporting gene is constructed, and these vectors are used to transfect into BEL7402 cells and HLE cells from human hepatocellular carcinoma HCC. As shown in Figure 17, the cells transfected with various vectors all show luciferase activity with various intensities,

indicating the transcription activities in these segments. The smallest fragment is a DNA segment (*pGL-PF4)* at about 38 bp in the upstream region of the transcription starting site. It possesses a basic promoting activity and functions as *LAPTM4* core promoter. The activity of *pGL-PF4* transfectant in BEL7402 is 20% of the reference promoter SV40, while the activity is low in HLE, only 6% of SV40 activity, about 1/3 of that in BEL7402. These data partially reflect the natural activity of *LAPTM4B* promoter in these two cell lines. It is consistent with the Northern blot results, where mRNA expression is high in BEL-7402 cell line and low in HLE cell line. Additionally, *pGL3-PF4* transfectant reveals dramatically different activities in these two cells. Its activity in BEL-7402 cells is 7 times higher than that in HLE cells. Apparently, different regulative mechanisms of *LATPM4B* gene transcription exist in BEL7402 and HLE cell lines.

In embodiments of this invention, the genome DNA is genotyped in order to determine the relationship between different *LATPM4B* genotypes and susceptibility of hepatocellular carcinoma. *LATPM4B* has two alleles, *LATPM4B*1* and *LATPM4B*2*, i.e., SEQ ID No: 6, is derived by PCR cloning. As shown in Figure 9, the difference between alleles **1* and **2* is the 19 bp sequence in the first exon 5' UTR. allele **1* has only one such sequence (nt 124~142dup, while **2* contains two such sequences in a tightly tandem arrangement (124-142dup, taking G at the transcription starting site TSS as +1 numbering standard). The insertion of the 19-bp sequence would eliminate the stop codon in 5'UTR in the corresponding **1* allele by a triplet shift. As a result, the open reading frame may be extended upstream by 53 amino acids at N terminus of the protein. The encoded protein by SEQ ID No: 6 then should contain 370 amino acid residues (SEQ ID No: 7). *LATPM4B* genotypes detected in human population are **1/*1, *1/*2* and **2/*2*, respectively (Figure 10). Studies show that the risk of developing hepatocellular carcinoma HCC for individuals with *LATPM4B* genotype **2/*2* is 2.89 times higher than individuals with non-**2/*2* type (Table 2). However, *LATPM4B* genotype in patients with esophagus carcinoma shows no difference from the normal population (Table 3). This indicates that *LATPM4B *2/*2* genotype correlates especially to susceptibility of hepatocellular carcinoma. As a result, *LATPM4B* allele *LATPM4B*2* provided by this invention can be used as a target to screen and diagnose people susceptible to hepatocellular carcinoma or having a high risk to develop hepatocellular carcinoma. Particularly,

using *LATPM4B *2/*2* genotype as a target to screen highly susceptible or high risk people can be more accurate.. **1/*1*, **1/*2* and **2/*2* of *LATPM4B* genotypes, *LATPM4B*2* encoded proteins or their antibodies, and/or *LATPM4B* extender and scavenger from human genome can all be used to screen people who are susceptible to hepatocellular carcinoma or having a high risk to develop hepatocellular carcinoma .

The expression vectors containing sequences described in SEQ ID No: 1, 2, 3, 6, 8, the transfection cell lines containing SEQ ID No: 1, 2, 3, 6, 8 sequences, and the primers amplifying SEQ ID No: 1, 2, 3, 6, 8 are all encompassed by this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1-A is the Northern Blot analysis profiles indicating the transcriptive expression of the gene of this invention in normal human liver, normal fetus liver and hepatocellular carcinoma tissues.

Figure 1-B is the Northern Blot analysis spectrum indicating the transcriptive expression of the gene of this invention in the human hepatocellular carcinoma cell lines.

Figure 1-C is a diagram showing the relationship of the expression level of the gene of this invention in human hepatocellular carcinoma tissues and the cancer differentiation status.

Figure 2-A is a diagram of hybridization in situ of hepatocellular carcinoma tissue. The *LATPM4B* mRNA in hepatocellular carcinoma nodule shows a strong positive staining.

Figure 2-B is an immunohistochemical diagram, where *LATPM4B* protein in hepatocellular carcinoma nodule shows a strong positive staining.

Figure 2-C is an immunocytochemical diagram, where *LATPM4B* protein is shown to exist in the transfected cells.

Figure 3 presents a Western Blot analysis diagram, where the expression spectra of *LATPM4B*-35 and *LATPM4B*-24 proteins encoded by the gene of this invention are shown in the tissues of normal liver (NL), hepatocellular carcinoma (HCC), and paired non-cancerous liver (PNL).

Figure 4 shows a growth curve of the accelerated proliferation of cDNA- transfected cells of this invention.

Figure 5 shows a column diagram, where the DNA synthesis of *LAPTM4B* cDNA-transfected cells of this invention is increased.

Figure 6 is a pie diagram showing an increase of cell numbers in S phase in cDNA-transfected cells of this invention (Flow cytometry analysis).

Figure 7 shows the oncogenic effect of cDNA- transfected cells of this invention on mouse.

Figure 8 is a histogram showing the level of the antigen of this invention in the serum of patients with hepatocellular carcinoma.

Figure 9 shows the partial sequence of *LAPTM4B* alleles of this invention.

Figure 10 shows the genotypes distribution of *LAPTM4B* alleles of this invention in human population.

Figure 11 is an immunohistochemical diagram of various cancer tissues derived from epithelium.

Figure 12 is a column diagram showing the inhibitory effect of LAPTM4B-EC2-pAb antibody on proliferation of hepatocellular carcinoma cells.

Figure 13-A, B, C, D, and E are the Western Blot analysis diagrams showing respectively that the expressions of cyclin D1, cyclin E, c-Myc, c-Fos, and c-Jun of cDNA- transfected cells of this invention are up-regulated.

Figure 14-A, B, and C are the analytical diagrams of the co-immuno precipitation, revealing respectively the interactions of the gene product (LAPTM4B protein) of this invention with α 6 β 1 integrin and with the epidermal growth factor receptor (EGFR).

Figure 15-A, B, and C are the analytical diagrams of the immunoprecipitation, showing the Tyr phosphorylation of LAPTM4B protein and the inhibitory effect of LAPTM4B-EC2-pAb on Tyr phosphorylation.

Figure 16-A and B are the analytical diagrams of the co-immuno precipitation showing respectively that LAPTM4B is involved in FAK-MAPK signal transduction pathway.

Figure 17 is a plot showing the transcriptive activity of various fragments of *LAPTM4B* promoter of this invention.

DETAILED DESCRIPTION OF THE INVENTION:

Sources of patients and normal control group:

57 cases of hepatocellular carcinoma patients, 50 males and 7 females, ranged in age

from 35-70. Their average age was 54±6.0. The tissues tested came from surgically excided specimens. The blood samples for the control group were collected from 206 similarly aged people with no symptoms and no cancer according to clinic tests and from 209 new born babies' umbilical veins.

109 esophagus cancer patients, 76 males and 33 females, ranged in age from 30-70. Their average age was 55±5.4. The test tissues came from surgically excided specimens. 116 people with no symptoms and no cancer, as determined by clinical examination, were selected as the control group S. Their blood samples were taken for testing. All the samples were extracted to obtain genomic DNA.

STATISTICAL METHOD

Chi-square (X^2) measurement and single factor ANOVA variance were used to treat and analyze the data.

Example 1: Northern Blot analysis of *LAPTM4B* expressions in four types of liver tissues at various proliferation and differentiation status.

Four types of liver tissues at various proliferation and differentiation status were chosen. They were from normal adult livers (NL, with very little proliferation and high differentiation), fetus livers (FL, at vigorous proliferation and low differentiation), hepatocellular carcinoma (HCC, out of controled proliferation and abnormal differentiation), and paired non-cancerous livers (PNL, generally is of precancerous stage in an active proliferation status). The Northern Blot analysis was used to detect the transcription of gene in these tissues. RNA samples were extracted from 5 normal adult liver tissues freshly obtained from surgical excision: 5 liver tissues from abortive fetus, 55 HCC tissues, and 55 paired non-cancerous liver tissues. After electrophoretic separation, they were transferred to a nylon film and hybridized by Dig labeled probe. The film was washed at 68°C and the hybridization signals were developed according to the manual. The results are shown in Figure 1. Band 1 was the sample from fetus livers. Band 2 was from the normal adult liver sample. Bands 3, 5, 7, and 9 were the samples from HCC. Bands 4, 6, 8, 10 were from PNL tissues. The results show that the expression of *LAPTM4B* in various liver tissues has the following order: HCC tissue > PPNL tissue and fetus liver tissue > normal adult liver tissue.

Example 2: Clonings of *LAPTM4B* gene, allele, and promoter

2-1 *LAPTM4B* gene cloning

By using fluorescence differential display technique, an unknown gene cDNA segment (LC27) was obtained from differential display spectum in four types of human liver tissues in different proliferation and differentiation status, such as normal adult livers (NL), fetus livers (FL), cancerous livers (HCC), and paired non-cancerous liver (PNL). The LC27 segment (426bp) was elongated by splicing homogenous sequences according to the EST to the 5' direction, and followed by RACE (rapid amplification of cDNA ends) and the high temperature RT-PCR techniques. Two full-length cDNA sequences, i.e., SEQ ID No. 2 and 3, were produced, and then confirmed by sequencing and BLAST program analysis.

2-2. *LAPTM4B* promoter cloning

The sequence of upstream region of the first exon of *LAPTM4B* gene at 5' terminal was obtained by biological informatics, and primers F1 and R1 were designed. Using human genomic DNA from HCC as the template, *LAPTM4B* promoter and the upstream sequence was obtained by PCR using Platinum Pfx DNA polymerase. After *Xho* I and *hind* III enzyme cutting, they were inserted into pGL3-Basicvector to form *pGL3-PF1*, and its sequence was determined (i.e., the test result see portion a of Figure 16).

As shown in Figure 17-A, no typical CCAAT (TTGCGCAAT) and TATA boxes were found in the *LAPTM4B* promoter sequence. In the upstream region of *LAPTM4B* promoter, there are manybinding sites for a variety of transcription factors, such as CREBP1/c-Jun, CEBP, PAX2/5/8, GATA, STAT, c-Ets-1, E2F, LYF-1, and c/v-Myb. They may function on regulation of *LAPTM4B* expression. In hepatocellular carcinoma, the abnormal expression and activation of these transcription factors possibly lead to an unbalanced expression of LAPTM4B proteins. Moreover, the *LAPTM4B* upstream region contains two highly homologous repeating sequences. It is worthwhile to further study on whether they have any effect on *LAPTM4B* expression regulation.

2-3. Cloning and sequence analysis of *LAPTM4B* alleles

2-3-1. DNA separation

Genome DNA was extracted from blood lymphocytes or cancer tissue samples from surgical excision of hepatocellular carcinoma and esophagus carcinoma according to the standard phenol-chloroform method.

2-3-2. Cloning and sequence analysis of the alleles

By using the same procedures for the promoter sequence cloning, two primers, F1: 5' GCGCTCGAGGCTCCAGGTG GAAGAGTGTGC 3' (inducing *Xho*I enzyme cutting site at 5' terminal sequence as indicated by underlining), and R1: 5' GCGAAGCTT GGACTTGGCCATGTGACCCG 3' (inducing *Xho*I enzyme cutting site at 5' terminal sequence as indicated by underlining), were designed and synthesized based on *LAPTM4B* gene sequence SEQ ID No. 3. The promoter sequence and its anterior sequence in the first exon of *LAPTM4B* were then cloned from human genomic DNA by PCR. The *pGL3-PF1* vectors constructed from various human genomic DNA were sequenced to screen the *LAPTM4B* alleles. The original *LAPTM4B* sequence was designated as *LAPTM4B*1*. The other one was designated as *LAPTM4B*2,* i.e., SEQ ID No. 6 in the sequence listings. Figure 9(A) shows the schematic diagrams of the *LAPTM4B* promoter and its first exon. The rectangle frame indicates the first exon, the black color area represents the encoding area, the white color is the non-coded area, and the gray area shows a 19bp DNA sequence. The horizontal line representing promoter part and F1, F2, R1, and R2 are where the four primers are located. "A" in the start codon ATG is defined as +1 in the sequence. Figure 9 (B) shows the partial sequences of the *LAPTM4B* alleles and their sequencing graphic spectra. The underlined part is a 19bp DNA sequence. The results reveal that *LAPTM4B*1* contains one copied 19bp DNA sequence and *LAPTM4B*2* has two copied 19bp DNA sequences, which are linked in the non-coded area (nt -33 — -15) of the first exon of *LAPTM4B*1*.

The sequence analyses indicate that *LAPTM4B*2* and *LAPTM4B*1* possess the same promoter. There is no difference in sequences between *LAPTM4B* alleles **1* and **2* promoters.

2-3-3. *LAPTM4B* genotype classification

E2 (5' GCCGACTAGGGGACTGGCGGA 3') and R2 (5' CGAGAGCTCCGAGCTTCTGCC 3') primers were designed and synthesized. A partial sequence of the first exon of *LAPTM4B* was amplified by PCR using templates of genomic DNA from normal people, hepatocellular carcinoma, and esophagus

carcinoma tissues. PCR conditions were as follows: 96°C pre-denature for 5 min; 94°Cfor 30 s, 68°Cfor 30 s, 72°Cfor 1 min, 35 cycles; 72°Cfor 5 min; then the PCR products were conducted to 2% Agarose gel electrophoresis analysis. Figure 10 shows *LAPTM4* gene **1/*1, *1/*2,* and **2/*2* three types in human population.

Example 3: Construction of the reporter plasmids and analysis of the promoter activity

A series of vectors, that contain the upstream sequences with various length of the *LAPTM4B* promoter, 5'UTR, the35bp encoding sequence in exon and the luciferase reporting gene, were constructed, i.e., the *LAPTM4B* gene promoter and the upstream sequence was cut by *Xho* I and I *Hind* III enzyme and connected to pGL3-Basic vector to form *pGL3-PF1,* and identified by sequencing. Then *pGL3-PF1* was used as a template, primers F2, F3, and F4 vs. R1 were used to amplify by PCR, respectively, to construct vectors, *pGL3-PF2, pGL3-PF3,* and *pGL3-PF4 which contain* promoter segments with various lengths and luciferase gene. These constructs were identified by sequencing.

The sequences of these primers are as follows:

F1: 5'GCGCTCGAG GCTCCAGGTGGA AGAGTGTGC 3 (nt –1341- -1321)
F2: 5' GCGCTCGAG TAA AAACGCTGTGCCAGGCGT 3' (nt –881- -861)
F3: 5' CCGCTCGAG TACCGGAAGCACAGCGAGGAT 3' (nt –558- -538)
F4: 5' GCGCTCGAG AGTAGAAGGGAAGAAAATCGC 3' (nt –38- -18)
R1: 5' GCGAAGCTT GGACTTGGCCATGTGACCCG 3' (nt 172-191)

These vectors were used to transfect BEL-7402 cells and HLE cells separately and the promoter activities were measured. As shown in Figure 17b, the vector-transfected cells all have luciferase activities with different intensities. *pGL3-PF3* showed similar activity in both BEL-7402 cells and HLE cells, which was about 27% of the SV40 promoter *(pGL3-Promoter)* activity. When comparing it with *pGL3-PF4* activity, however, there was almost no difference in BEL-7402 cells. In HLE cells, *pGL3-PF3* activity was 7 times higher than *pGL3-PF4.* As shown in Figure 17a, *pGL3-PF3* (—41~ —558) has many potential binding sites for transcription factors. One or many of them, especially c-Ets-1, may play a regulating role in HLE cells and make the luciferase activities of *pGL3-PF3* and *pGL3-PF4* tranfectants remarkably

different in HLE cells. The *pGL3-PF3* activity is higher than that of *pGL3-PF1* and *pGL3-PF2* in both BEL-7402 and HLE cells, implying that some negatively regulatory factor exists. One or more of these negatively regulatory factors bind with the promoter upstream target sequence (—558 upstream) to induce a downregulated *LAPTM4B* gene expression. This suppressive effect was stronger in HLE cells than in 7402 cells. This means that HLE cells may contain some factors that strongly suppress the expression of *LAPTM4B*. The Northern Blot analysis presented in Figure 1-B also shows a low expression of *LAPTM4B* in HLE cells, supporting the above hypothesis. The *pGL3-PF2* vector contains two DNA repeating fragments (—41~ —328, —574 ~ —859), which is one more DNA fragment (—574 ~ —859) than *pGL3-PF3*. *pGL3-PF3* exhibits higher activity than *pGL3-PF2* in both cells. This result indicates that the two repeating sequences negatively regulate gene transcription. They have many potential binding sequences for the transcription factors and provide two binding sites for each negatively regulating factor. Since many transcription factors often form dimers, they have to bind with two target sequences to be able to function. In the case of *pGL3-PF3,* which can only provide one binding site, no function is shown. Since the *pGL3-PF3* transfectant has a disinhibitory effect, its activity is higher than other vector transfectants.

Example 4: Western Blot analysis of LAPTM4B protein expression

The tissue sample was placed on ice and cut into small pieces by scissors. 0.1 gram of wet tissue was selected and placed in a manually operated homogenizer. 1 mL lysis buffer was added in each tube and the mixture was thoroughly homogenized. The lysate was transferred to a fresh tube and centrifuged at 4°C, 12 000 g for 10 min to remove the debris. If cells are used, the cells in a culture dish were digested with 0.25% tripsin buffer, followed by two PBS rinses and centrifuged at 500g for 3 min. The cleared supernatant was collected, and the proteins in the supernatant were separated by SDS-PAGE electrophoresis, and then transferred to the NC membrane. The membrane was blocked at 4°C overnight with 5% non-fat powdered milk in a TBS buffer. containing, 0.05% Tween 20 Then it was incubated with the rabbit polyclonal antibody, LAPTM4B-EC2 $_{232\text{-}241}$-pAb (1:500 dilution) or mouse Anti-FLAG M2 monoclone antibody (Sigma, 1:750 dilution) at room temperature for 2 hours, and then rinsed with TBS for three times. It was further incubated with a peroxidase-coupled second antibody (IgG), such as goat anti rabbit or goat anti mouse

(1:3000 dilution), for 2 hrs, followed by three rinses with a TBS buffer (pH 8.0, containing 0.05% Tween 20). The last wash was with a buffer containing no Tween 20. ECL (Santa Cruz) was used to expose the positive bands (performed as manufacturer's instructions). When two antibodies were sequencially hybridized in one membrane, the ECL exposed membrane was rinsed first with TBS followed by washing with 30 mL TBS (2% SDS and 210 μL β-mercaptoethanol) for 30 min at room temperature. The 30 min TBS rinse removed the previous antibody and its signal in the membrane, which then could be used for the second hybridization. Figure 3 shows that LAPTM4B-35 was over expressed in HCC tissues and HCC cell lines.

Example 5: Regulatory effect of the gene of this invention on cell proliferation and the malignant phenotype of cancer cells as demonstrated by a full-length cDNA transfection.

Using *pGEMT-E2E7* plasmid as a template and the PCR method, a full length or partial cDNA, or the reading frame of LAPTM4B gene was amplified by PCR with primers A, or B and E, and the Pfx DNA polymerase. *Bam*HI enzyme cutting site (GGATCC) and ribosome binding site sequences (GCCACC) were introduced in primer A and B at 5' terminal and *Eco*RI enzyme cutting site (GAATTC) was incorporated in the primer E. The amplified products AE and BE were digested by restriction enzymes *Bam*HI and *Eco*RI, purified, and ligated into *pcDNA 3.0* vector. They were transformed conventionally to DH5 E. coli and the positive clone was selected, and the constructed plasmid was sequenced for identifying. The constructed plasmids were named as *pcDNA3/AE* and *–BE,* respectively. *pcDNA3/AE* contains a full-length ORF, while *pcDNA3/BE* contains the ORF starting from the second ATG to TAA. Compared with *pcDNA3/AE*, *pcDNA3/BE*- encoded protein is missing 91 amino acids at the N terminal.

Mouse BHK, NIH3T3 cell lines and human hepatocellular carcinoma HLE cell lines, in which the expression of *LAPTM4B were* all at very low level, were transfected by *pcDNA3/AE* or *–BE,* and clones that *LAPTM4B* expression were stable and high were selected. The total viable cell numbers were determined by the acidic phosphatase method and the cell growth curve was plotted. The cell cycle was analyzed by the

flow cytometry. The expression levels of cell cycle-regulating protein, including cyclin D1 and Cyclin E, and proto-oncogene products, including c-Myc, c-Fos, and c-Jun (transcription factors for regulating cell proliferation) were measured by the Western Blot analysis. The results show that the cell proliferation was accelerated after being transfected by *LAPTM4B-AE* expressive plasmid (Figure 4, 5, 6). Expressions of cyclin D1, cyclin E, c-Myc, c-Fos, and c-Jun were also greatly increased (Figure 13-A, B, C, D, E, respectively). But the dependence of growth on serum in LAPTM4B-35-overexpressed cells was greatly reduced (HLE-AE cell proliferation proceeded normally in 1% FCS, but HLE and HLE-MOCK cells stop proliferation at the same condition). In the meantime, the ankorage-dependent cell growth of HLE-AE cells was clearly weakened. Large colonies of HLE-AE cells were formed in the soft agar gel, which indicates that this gene participated in the regulation of cell proliferation and its over expression (activation) was related to the dysregulation of cell proliferation. Furthermore, the migrating capability of HLE-AE cells was also enhanced (The HLE-AE cells that migrated through the membrane pores were increased from 1216±403.8 for the control to 4082.5±748.8). Its capability to invade Matrigel was also greatly increased (from 25±12.73 cells for the control to 1325±424.26 cells). The results show that LAPTM4B over expression promotes the development of cell malignant phenotype. On the contrary, BHK-BE, NIH3T3-BE, and HLE-BE cells transfected by *LAPTM4B-BE* expressive plasmid could not form clones. They were all dead within three weeks. These results demonstrate that LAPTM4B-24 plays antagonistic roles to LAPTM4B-35.

Example 6: Tumorigenic effect of *LAPTM4B cDNA*- transfected cells on mouse

Six-week old male mice were randomly selected and divided into three groups: In the first control group, the mice were injected with physiological saline. In the second control group, the mice were inoculated with the *pcDNA3* MOCK (no-load plasmid) transfected cells by. In the test group, all the mice were inoculated with *pcDNA3 /AE* (a plasmid containing full-length cDNA) transfected NIH3T3 cells. Each mouse was subcutaneously inoculated with $2x10^6$ cells. There were four to six mice in each group. The mice were sacrificed after 21 days inoculation and dissected. As shown in Figure 7, two mice (half of inoculated mice) in the test group developed a clearly moderate malignant fibrosarcoma (A, B); the other two mice were identified as lymphatic tissue

at the inoculated sites (C, D). In contrast, twelve mice in the two control groups showed no sign of tumor formation after being inoculated for 86 days.

The results in Examples 4, 5, and 6 indicate that *LAPTM4B* may be a novel proto-oncogene.

Example 7: Primary analysis of LAPTM4B antigenin the serum of patients with hepatocellular carcinoma by the ELISA method

96 wells culture plates were coated with sera in various dilutions from HCC patients and normal people by known agreement at 4°C overnight. Each well was washed with 0.5% Tween-20 washing solution, and then 2% BSA was added for blocking at room temperature for 1 hour. Then LAPTM4B-EC2-pAb antibody in various dilutions was added and incubated for 2 hours at room temperature. The goat anti-rabbit antibody labeled by horseradish peroxidase (1:1000 times dilution) was added after PBS washing. After standing at room temperature for 2 hours and one PBS washing, 1 g/mL o-phenyldiamine was added for 10-15 minutes to develop color and H_2SO_4 was used to stop the reaction. The microtiter for enzyme analysis was used to measure OD. at 490 nm and the antigen level was estimated. The results are shown in Figure 8. Clearly, the sera of patients with hepatocellular carcinoma contained higher level of LAPTM4B antigen than that from normal people, indicating that LAPTM4B has a potential to become a new marker for hepatocellular carcinoma diagnosis.

Example 8: Functional determination of LAPTM4B protein in signal transduction by co-immunoprecipitation o and antibody inhibition analysis.

The cell lysate was prepared according to the method in Example 4. The first antibody was added to the supernatant. After 1 hour's shaking at 4°C, 50 μL protein G-Agarose suspension was added and the mixture was shaken at 4°C for at least three hours or overnight. The immunocomplex precipitate was collected after centrifuging at 12000g for 20 seconds. The complex was re-suspended by adding 1 mL washing buffer I and shaken at 4°C for 20 min. The mixture was centrifuged at 12000g for 20 seconds and the supernatant was removed carefully. This step was repeated once. Then the complex was re-suspended by adding washing buffer II, shaken at 4°C fro 20 min.,

and centrifuged at 12000g for 20 seconds. The supernatant was removed carefully. The last two steps were repeated once. The complex was re-suspended by adding washing buffer III, shaken at 4°C fro 20 min, and followed by 12 000 g centrifugation for 20 seconds. The supernatant was removed completely. 50 μL 1xSDS loading buffer was added in the precipitate and the mixture was boiling in 100°C water bath for 5 min to denature and dissociate the immunocomplex in the sample. After 12000 g centrifugation for 20 second, the supernatant was removed and analyzed in SDS-PAGE apparatus.

BEL-7402 cell was preincubated for 0, 10, 20, and 40 min, respectively, on LN-1 substance in serum free medium. Co-immunoprecipitation was performed with LAPTM4B-EC2-pAb o from the cell lysate. The co-immunoprecipitates were respectively adsorbed by Protein G-Sephorose, centrifuged, and analyzed by 10% non-reductive SDS-PAGE. Then the phosphorylations of LAPTM4B, FAK and MAPK were analyzed separately by the Western Blot with p-Tyr mAb.

BEL-7402 cells were preinoculated separately with LAPTM4B-EC2-pAb (15 μg/mL) and anti-Glut2 (15 μL/mL) antibodies at 37°C under 5% CO_2 for 2 hrs, and then seeded on LN-1 substance and incubated for indicating time. . Under the same conditions, the anti-Glut2 antibody treated cells and no antibody treated cells were used as control. The cell lysate in each group was analyzed by the Western Blot analysis with p-Tyr mAb. The inhibitory effects of various antibodies on phosphorylation of LAPTM4B were analyzedThe results show that LAPTM4B-35 was phosphorylated peakly when human hepatocellular carcinoma BEL-7402 cells were attached on laminin substrate. The phosphorylation of LAPTM4B-35 reached the highest level in 10 min after cell attachment (Figure 15-A). Meanwhile LAPTM4B-EC2-pAb could inhibit almost completely its phosphorylation (Figure 15-B), while the anti-Glut2 (an antibody against a non-related plasma membrane protein Glut2) showed no such inhibitory effect (Figure 15-C). On contrary, LAPTM4B-24 cannot be phosphorylated. The phosphorylation of LAPTM4B-35 Tyr_{285} would form a binding site for signal molecules that contain SH2 domain. In the meantime, LAPTM4B-35 itself presents typical binding sites for signal molecules that contain SH3 domain. Therefore, LAPTM4B-35 functions most likely as a very important docking protein of molecules for signal transduction or a special organizer of

membrane microdomain. It could recruit signal molecules related inside or outside cells, so that to play pivotal roles in signal transduction associated with cell proliferation, differentiation and apoptosis. Moreover, the attachment of human hepatocellular carcinoma cells on laminin substrate can also cause Tyrphosphorylation of the cytoplasmic signal molecule FAK (Figure 16-A),and the LAPTM4B-EC2-pAb and anti-integron α 6 mAb against the epitope of the extracellular region of α6 both can prevent FAK phosphorylation without affecting the expression level of FAK protein by preincubating with BEL 7402 cells.-Similarly, the attachment of BEL 7402 cells on laminin substrate can also induce Tyr phosphorylation of the signal molecule MAPK (Figure 16-B), and its phosphorylation can be inhibited by preincubating cells with LAPTM4B-EC2-pAb without changing the expression level of MAPK protein. These results indicate that the interaction between LAPTM4B-EC2 domain (the second extracellular region) and integrin α 6 subunit plays an important role in triggering FAK-MAPK signaling pathway.

The results from Examples 4-8 suggest that LAPTM4B-35 can be potential targets of drugs for regulating cell proliferation, differentiation, and apoptosis.

Example 9: *LAPTM4B* genotype classification

LAPTM4B genotypes in genomic DNA from blood of normal individuals and patients with hepatocellular carcinoma were detected by PCR. Two primers were designed and synthesized according to the flanking sequence of 19bp DNA sequence in *LAPTM4B* gene sequence 3:

F2: 5' GCCGACTAGGGGACTGGCGGA 3'

R2: 5' CGAGAGCTCCGAGCTTCTGCC 3'

The partial sequence of the first exon was amplified using genomic DNA as a template. PCR conditions were as follows: 96°C pre-denature for 5 min, 94°Cfor 30 sec, 68°Cfor 30 sec, 72°Cfor 1 min, 35 cycles, 72°C extension for 5 min. PCR products were analyzed by 2% agarose gel electrophoresis and the results are shown in Figure 10. The lanes 1, 6, 12, and 13 represent a 204 bp nucleotide segment in *LAPTM4B*1/*1*. The lanes 5, 8, 9, 14, and 15 represent a 223 bp nucleotide segment in *LAPTM4B*2/*2*. The lanes 2, 3, 4, 7, 10, and 11 represent 204 bp and 223 bp nucleotide segments in *LAPTM4B*1/*2*. Line M is the marker. The results reveals

that in the homozygous gene pair of **1/*1* or **2/*2* either the 204bp or 223bp DNA segment was amplified, while in **1/*2* hybrid gene pair 204bp and 223bp DNA segments were both amplified simultaneously. Therefore, the genotype of *LAPTM4B* in Chinese population can be classified as *LAPTM4B*1/*1, *1/*2,* and **2/*2* (Figure 10).

Example 10: Frequency distribution of *LAPTM4B* genotypes and alleles in normal people and patients with hepatocellular carcinoma

In one of the embodiments of the present invention, the occurrence frequency of *LAPTM4B* genotypes in 209 normal Chinese and 57 patients with hepatocellular carcinoma was analyzed and compared in Table 2. The Hardy-Weinberg equation was used to get the expectancy analysis. The frequency of *LAPTM4B* allele **1* and **2* from patients with hepatocellular carcinoma differs significantly from that of normal people. Their ratios are 0.5175: 0.6746 and 0.4825: 0.3254, respectively. The occurrence frequencies of *LAPTM4B* allele **1* and **2* in a normal population are 0.6746 and 0.3253, while the occurrence frequency of *LAPTM4B* allele **1* and **2* in patients with hepatocellular carcinoma are 0.5175 and 0.4825. The occurrence frequency of genotype **1/*1* (p=0.029) and **2/*2* (p=0.003) in the group of hepatocellular carcinoma patient shows a significant statistical difference from its control group. In the hepatocellular carcinoma patient group, only 29.8% is of **1/*1*, while in the normal control group, 45.93% is of **1/*1*. The occurrence frequency of **2/*2* genotype in the hepatocellular carcinoma patient group is 26.32% as compared to 11.01% in the control group, therefore its occurrence frequency is increased significantly ($p < 0.01$). The analysis shows that the risk suffering from HCC of individuals in **2/*2* genotype of is 2.89 times greater than that in other genotype in developing hepatocellular carcinoma. Thus, the *LAPTM4B*2/*2* genotype is correlated with the susceptibility of developing hepatocellular carcinoma.

As shown in Table 3, patients with different LAPTM4B genotypes did not show any differences in hepatocellular carcinoma Grade, stage, or HBV infection. 83.3% of the HCC patients have a positive HBV.

Table 2: Distribution of *LAPTM4B* genotype in hepatocellular carcinoma patients and normal population

	N (%)		*P* Value
	Control B (n=209)	Hepatocellular carcinoma group (n=57)	
LAPTM4B genotype			
**1/*1*	96 (45.93)	17(29.82)	0.029 [a]
**1/*2*	90 (43.06)	25 (43.86)	0.914
**2/*2*	23 (11.01)	15 (26.32)	0.003 [b]
Frequency of alleles			
**1*	0.6746	0.5175	
**2*	0.3254	0.4825	

[a] OR: 0.500, 95%CI: 0.267 – 0.939; [b] OR: 2.888, 95%CI:1.390 – 6.003 (OR risk suffering HCC, and 95%CI is confidence interval)

Table 3: Clinical data of the hepatocellular carcinoma patients used in *LAPTM4B* genotype classification

	LAPTM4B Genotype			*P* Value
	**1/*1*	**1/*2*	**2/*2*	
Total number	17	25	15	
Males	14	24	12	
Females	3	1	3	NS
Cancer Grade G1				
G2	0	2	0	
G3	1	4	8	
G4	7	7	4	
Cancer stage	9	12	3	NS
I				
II	0	0	0	
III	5	8	5	
IV	4	7	3	
HBV Infection	8	10	7	NS
Negative				

Positive	1	4	4	
No diagnosis	13	16	10	
	3	5	1	NS

NS: No significant difference

Example 11: Frequencies of genotype and allele in patients with esophagus carcinoma

To study if the *LAPTM4B* genotype is related to the susceptibility of developing other cancers, the genomic DNA from blood of 116 normal people and 109 patients with esophagus carcinoma from the same location were analyzed. As shown in Table 4, *LAPTM4B* genotype of patients with esophagus carcinoma is no significant different from control group of the normal population. *LAPTM4B* alleles are not related with the susceptibility of developing esophagus cancer.

Table 4: Distributions of *LAPTM4B* genotypes of patients with esophagus carcinoma and normal population

	N (%)			*P* Value
	Control group B (n=209)	Control group S (n=116)	Esophagus carcinoma (n=109)	
LAPTM4B genotype				
**1/*1*	96 (45.93)	52 (44.83)	49 (44.95)	>0.05
**1/*2*	90 (43.06)	49 (42.24)	48 (44.04)	>0.05
**2/*2*	23 (11.01)	15 (12.93)	12 (11.01)	>0.05
Allele frequency				
**1*	0.6746	0.6595	0.6697	
**2*	0.3254	0.3405	0.3303	

Example 12: LAPTM4B-35 expression in some epithelium sourced cancers

The relationship between the LAPTM4B-35 protein expression and other cancers was

studied by an immunohistochemical method. The fixed specimens from esophagus cancer, breast cancer, lung cancer, stomach cancer, colon cancer, and rectal cancer positive tissues and the negative control noncancerous tissues were obtained from surgical excision and treated according to the following steps:

1. Specimen dewaxing by xylene
2. Katocromy with different concentrations of ethanol, 100% — 95% — 90% — 80% — 70%. PBS was used to remove endogenous peroxidase
3. Antigen repairing by sodium citrate
4. PBS rinse twice
5. Normal goat serum blocking
6. Keep LAPTM4B-$N_{1\text{-}99}$pAb at 37°C for 1 hour
7. PBS rinse three times
8. Keep HRP labeled goat anti- rabbit antibody at 37°C for 1 hour
9. PBS rinse three times
10. Develop color by DAB
11. Nuclear retaining with hematoxylin
12. Ascending dehydration by ethanol at different concentrations (70% — 80% — 90% — 95% — 100%
13. Mounting

As shown in Figure 11, The 11-A indicates a normal esophagus tissue (Negative), B is an esophagus cancer tissue (Positive), C is a normal breast tissue (Negative), D is the breast cancer tissue (Positive), E is a normal lung tissue (Negative), F is a lung cancer tissue (Positive), G is a normal stomach tissue (Negative), and H is a stomach cancer tissue (Positive). As can be seen from the figures, LAPTM4B was clearly expressed in lung cancer, stomach cancer, and breast cancer tissues, while it was not expressed clearly in esophagus cancer and large intestine cancer.

INDUSTRIAL APPLICATIONS

The proteins encoded by *LAPTM4B* gene in this invention could be possibly used as new markers in early diagnosis of some cancers. By using the widely applied ELISA method in clinical tests, and the prepared related testing reagent kits, the efficiency and the accuracy of the early diagnosis of cancers, especially the primary hepatocellular carcinoma, can be improved.

LAPTM4B gene can be used as target gene in the cancer treatment. suppressing LAPTM4B-35 expression and promoting LAPTM4B-24 expression could inhibit the growth of hepatocellular carcinoma cells, reverse malignancy phenotype or delay its development. For example, the expression products of *LAPTM4B* gene, LAPTM4B-35 could be inhibited by the newly developed siRNA interference technology. Furthermore, *LAPTM4B*-BE-cDNA could be recombinated in the engineered virus expression vector and be used in antitumor gene therapy through an up-regulation of LAPTM4B-24 expression. LAPTM4B-35 protein could also be used as a new target for pharmaceutical treatment. Since LAPTM4B-35 protein can function as an assembling platform for complex of cell signal transduction molecules, and it contains a number of binding sites for signal molecules, there is a great potential to develop various new medicines with LAPTM4B protein as targets. Moreover, this invention has initially demonstrated that LAPTM4B-EC2-pAb antibody can inhibit tumor cell proliferation and block its signal transduction. Based on the discovery in this invention, further studies can be pursued on the possibility of using antibody to inhibit hepatocellular carcinoma and some other cancer development. After a better understanding on the effect, a humanized soluble single chain antibody could be developed for clinical treatment on HCC patients. Peptide vaccines could also be developed. If the vaccines can be successfully made, it will not only help cure hepatocellular carcinoma and some other cancer, but also prevent cancerogenesis in the high risk population. In summary, many new anticancer approaches can be developed based on the embodiments of this invention. As important supplements for treatments of hepatocellular carcinoma and other cancers, this invention will help increase the cure rates of hepatocellular carcinoma and other cancers. This project would generate a significantly great impact on human society.

In specific embodiments, *LAPTM4B* genotype of genomice DNA is genotyped. The relationship of various genotypes with the susceptibility to hepatocellular carcinoma as well as with other cancers s is investigated. It is discovered that one of the genotypes, *LAPTM4B*2/*2,* is correlated closely to hepatocellular carcinoma susceptibility. As a result, it provides a new and accurate criterion for screening people who are susceptible to primary hepatocellular carcinoma in the high risk population. It is of important significance to the assessment and prevention of high risk population from developing hepatocellular carcinoma.